

Evaluating urban public facilities of Shenzhen by application of open source data

Miaoxi Zhao, Gaofeng Xu & Yun Li

To cite this article: Miaoxi Zhao, Gaofeng Xu & Yun Li (2016) Evaluating urban public facilities of Shenzhen by application of open source data, Geo-spatial Information Science, 19:2, 129-139, DOI: [10.1080/10095020.2016.1176724](https://doi.org/10.1080/10095020.2016.1176724)

To link to this article: <http://dx.doi.org/10.1080/10095020.2016.1176724>



© 2016 Wuhan University. Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 17 May 2016.



Submit your article to this journal [↗](#)



Article views: 393



View related articles [↗](#)



View Crossmark data [↗](#)

Evaluating urban public facilities of Shenzhen by application of open source data

Miaoxi Zhao^{a,b}, Gaofeng Xu^a and Yun Li^c

^aDepartment of Urban Planning, South China University of Technology, Guangzhou, China; ^bDepartment of Geography, Ghent University, Ghent, Belgium; ^cDepartment of Urban Planning, Shenzhen University, Shenzhen, China

ABSTRACT

This article applies open source data of public facilities through data mining, not only to evaluate the public facilities from an objective dimension, but also to reflect the sensory opinions of the group factually, eventually realizing the evaluation measurement of urban public facilities. The research takes Shenzhen city as an empirical case and chooses typical public facilities to mine data, resolve address and weight to explore the application of public facilities evaluation under dimension reduction of open source data. The empirical study consists of three parts. First, as the objective evaluation, we estimate the density distribution and per capita of public facility through data mining and address resolution. Second, as the subjective evaluation, we carry on the location analysis to high-score public facility through attention and satisfaction data of Internet evaluation. Finally, as mentioned above, we calculate the weight of objective and subjective evaluation of public facility, eventually formatting the comprehensive evaluation of public facilities.

ARTICLE HISTORY

Received 17 January 2016
Accepted 25 March 2016

KEYWORDS

Open source data; big data;
public facility; Shenzhen

1. Introduction

Since the new century, the mega-city with more than 5 million of the population is emerging all over the world, and these mega-cities have brought a lot of convenience for people living in. But at the same time, it has also caused a lot of problems. Public facilities become one of the main objects of criticism. As reported by Hu and Wu (2011), with Chinese urbanization in focus for a long time, Sir Peter Hall mentioned that, “Chinese cities have faced enormous challenges for its citizens in a decent quality of life”.

In reality, it has been difficult to achieve mega-city public facilities in overall scale in the field of urban studies, which more rely on the way of questionnaire and interview. Although this small sample of research method could discriminate pros and cons of single building or micro-built environment, it always failed in the scale of the whole city, which means that it cannot recognize the quality features of city public facilities from the macro perspective. “See the wood but missing forest”, the way of interview cannot objectively reflect the real quality of public facilities (Delyser and Sui 2013). On the other hand, the evaluation of urban public facilities cannot be separated from the individual’s feelings and demands. With the help of statistical data, the quality of urban public facilities at overall scale could be calculated. However, this kind of structural measurement

usually ignores the true feelings of the individual, and even forms a clear fracture with the daily life recognition.

With the continuous development of information technology, the Internet data have not only objectively recorded all kinds of spatial information, but also performed the subjective evaluation of public places in the city. Previous study on urban computing and big data has mostly been explored by interdisciplinary studies in the field of computer professionals; however, it is often difficult for them to take into consideration of urban planning. This article takes exploratory improvement based on existing technology application. Specifically, the research tries to obtain open source data of public facilities through data mining, not only to evaluate the public facilities from objective dimension, but also to reflect the sensory opinions of the group factually, eventually realizing the evaluation measurement of public facilities.

2. Research overview

On the study of public facilities evaluation, scholars used objective data by various models, mainly focusing on assessment and measurement of the accessibility and fairness (Chen 2012; Han and Lu 2012; Yang and Xu 2015). In other studies, the ecological and economic factors were introduced into the space evaluation of public facilities, and the suitability of the layout of urban

public facilities under the influence of multiple factors was discussed (Yang, Shi, and Deng 2010; Sun, Wang, and Yao 2015). For instance, Zhang, Wei, and Hua (2012) used the location quotient index to calculate the spatial distribution of public facilities, and then, from the perspective of spatial distribution of public facilities, discussed the overall service utility evaluation problem. But with the rapid growth of population, urban spatial structure, especially the spatial structure of mega-cities is increasingly diverse. Therefore, to the demand for public facilities as results, space distribution has become a part of the overall service utility evaluation (Gu and Yin 2010). This kind of evaluation method is easy to fall into the problem of “see the wood but missing forest”, ignoring the real-life experience of any individual resident. Some scholars evaluate the public facilities based on difference of age structure, occupation status, and use purpose from perspective of the users’ demands (Zhao 2009; Ren 2014), but the heteronomy factor of such a small-sample survey method cannot achieve the performance evaluation of city spatial structure (Zhang 2014). At the present stage, there is lack of research which combines subjective and objective approach to evaluate public facilities in the whole scale of city.

As Delyser & Sui (2013) and Zhang (2014) reported that in recent years, the emergence of the Internet open source data to a certain extent solves the problem of the data shortage. This makes accessible some unfeasible quantitative analysis. In the measurement of urban activities, many scholars acquired data from multiple ways such as telecom operators, social networking sites, taxi and bus intelligent card, etc. (Becker et al. 2011; Mark and Nick 2011; Long, Zhang, and Cui 2012; Naaman, Zhang, and Brody 2012; Batty 2013; Kang et al. 2013). Gui, Xiang, and Li (2012) presented a parallel algorithm based on MapReduce to recognize city hot pot. In terms of urban spatial structure, Hollenstein and Purves (2010) detected the central area of London and Chicago metropolitan area by 8 million pictures with locational information from Flickr. Based on distinguishing the urban spatial geographical features by the terrain map database, Lüscher and Weibel (2013) took major cities in the UK as the case and realized the division of urban central area with the resident emotion data, as Cranshaw et al. (2012) and Qin, Zhen, and Xiong (2013) reported. Wang, Zhu, and Xie (2014) used mobile signaling data to explore the hinterlands of Wujiaochang, Daning. Niu, Ding, and Song (2014) used mobile signaling data for application of urban spatial structure analysis. Liu, Fang, and Guo (2014) especially systematically summarized the studies on people’s behavior patterns and geographic situation based on location information. Yu and Ai (2015) took public facilities as an example, using Kernel density method for density distribution analysis and visualization, put forward advantages to the expression methods of the Venn diagram and others. The emergence of these

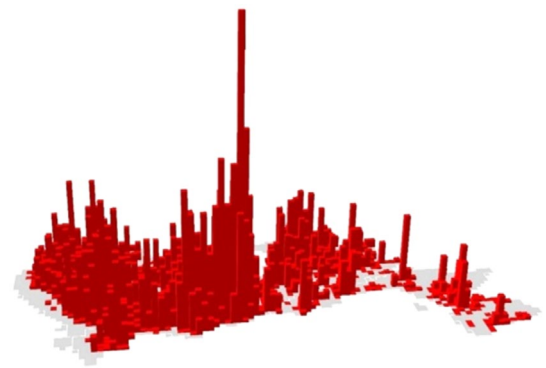


Figure 1. Density distribution of Shenzhen public service facility in www.dianping.com.

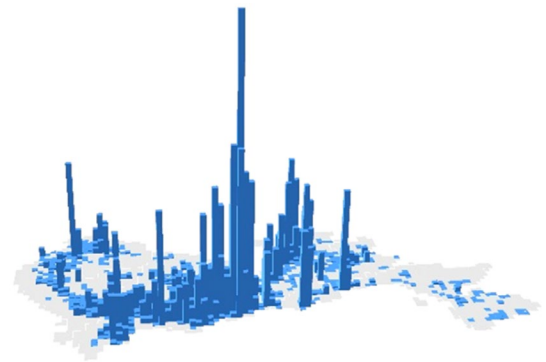


Figure 2. Comment density distribution of Shenzhen public service facility in www.dianping.com.

relevant research results provides the technical support for the development of urban planning in the new era.

As Castells (1996) and Kitchin (2013) reported, information technology has accelerated the time and spatial exchange of knowledge, technology, talent, capital, etc. “Space of flow” becomes the main carrier of resident activities, and reflects the urban space through a large number of Internet or mobile information equipment data form. Thus, in the process of production and consumption of the Internet, it is forming a wealth of evaluation data (Figures 1 and 2), and these address information has become an important source of information for public facilities evaluation.

3. Research design

3.1. Data acquisition

In the course of the research, the main data source is from the Web www.dianping.com, and we take 2014 enterprise list as the data complement. Through the data mining of subjective/objective evaluation and spatial location on the Internet, we then explore the structural characteristics of urban public facilities evaluation. The acquisition time of Internet data is July 2014. Public facilities are acquainted from the list of enterprises and institutions from Trevor database. For Internet data, we use Python software to improve the data access tool,

then automatically output the facility name, the type, the number of evaluation, the evaluation score, and the address information. After the completion of open source data collection, we geocode address data through the Baidu open platform. Specifically, through the open platform of Baidu LBS Place Service API Web and API Geocoding Service functions, we get access to data information of various types of services. The basic steps are as follows: (1) Get access to the API interface key; (2) Use the key, and spell HTTP to request URL according to the place area to retrieve common interface parameters; and (3) Based on the Locoy Spider software platform, the data acquisition field information is written and the data are sent and received by the HTTP request. Figures 1 and 2 that show the spatial distribution of public comment network data are the examples of the result of data mining process.

To achieve the research object, we select the city of Shenzhen because it is an economically developed city and the size of its population also ensures that the data could be applied to the required space scale. Aiming at public facilities evaluation, the subjective/objective dimension is divided, and the distribution information of public facilities and service population information are the objective measure data. The evaluation of the facilities in the site is a subjective measure data. In addition, the research complements the school and nursing home facilities of Shenzhen city from the list of enterprises and institutions in 2013, and dummy variable assignment of 0/1 is used to distinguish the key middle and primary school.

From a holistic perspective, the data have the multidimensional characteristic of big data. For example, using the data from www.dianping.com, enterprises and institutions have millions of items about facility information, of which the data format is diversified. As a result, the effective information of each database for the case city is sparse. The value of the data-set is eventually low as well. Specifically, a relatively detailed evaluation is displayed on www.dianping.com, but the database facilities of schools and nursing homes in enterprise list are not part of this content. In order to avoid such a situation, the data need to be identified by other auxiliary information. In addition, facilities address cannot directly correspond to consumer groups. The study takes the grid method to merge the data in accordance with the rules of the space grid unit. This method also helps to reduce the amount of calculation work.

3.2. Research method

In order to make the facility data more convenient for spatial analysis, the research transforms the spatial data into 1-km × 1-km grids for conventional software can be analyzed. This is the key step for the comprehensive evaluation of the public facilities in the mega cities. In order to grasp the overall features of the urban public

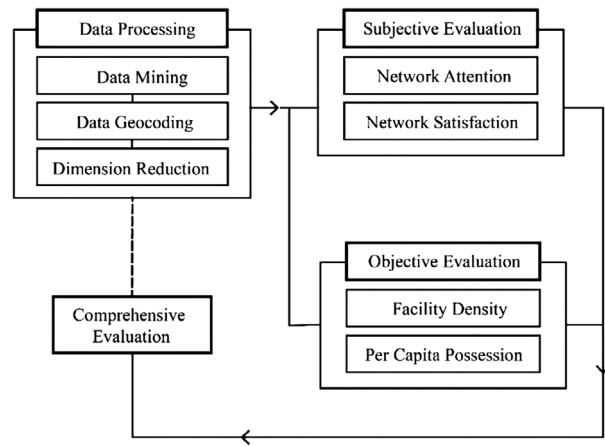


Figure 3. Methodology on public service facilities evaluation.

facilities, the spatial grid dimension of the open source data can be matched with the census data of the GIS platform under the premise of the accuracy. On the basis above, according to the mixed characteristics of data, the study uses the evaluation system to include the objective evaluation, subjective evaluation, and through the comprehensive evaluation of urban public facilities to achieve the measurement of location difference. The specific research includes the following three aspects as shown in Figure 3.

First, in the objective evaluation of urban public facilities, the distribution density of the facilities and the per capita possession is measured. Specifically, we use the Web tool www.dianping.com, enterprise list, and other multidimensional data to geocode address, and then get access to the location of various facilities. Combining with the spatial distribution of urban construction land and population, we further analyze distribution density of different locations of various facilities in the city, per capita possession, and other secondary objective indicators for the overall structure of the objective evaluation.

Second, in the subjective evaluation of urban public facilities, the spatial distribution of positive evaluation value is further distinguished based on the calculation of the average evaluation value. After geocoding the network evaluation data, the spatial distribution characteristics of the attention rate (the number of comments) and the degree of satisfaction (positive evaluation) of secondary subjective index, based on the comments data from www.dianping.com, are analyzed, which is for the subjective evaluation of individual perception dimension.

Finally, we weight the subjective/objective evaluation of the public facilities to explore the comprehensive evaluation of urban public facilities. In specific calculation, we first obtain all kinds of secondary index value. Then we determine the index weight according to the variance of the value in the basic elements of each facilities type. Third, the comprehensive weights are determined according to the variance and Delphi method. Finally, facilities evaluation indexes are thus obtained.

Of course, our main purpose is to explore the application of open source data for space research. The selection of the typical elements does not affect the exploration of the basic types of research.

4. Objective evaluation of public facilities

The objective aspect of the facility evaluation takes into account of the relevant specifications of the planning and design, including “Urban Residential Area Planning and Design Standards 50180-93 GB (2002)”, “Guidelines of Shanghai Residential Area Public Facilities DGJ08-55-2006-2”, “Shenzhen Urban Planning Standards and Guidelines (2004)”. Those criteria show that the rationality of the planning of public facilities generally consists of two dimensions: the service radius and the thousand-person index. In this study, the density distribution of each type of public facilities directly reflects the service facilities radius. The per capita amount is another way of reporting thousand-person index. From the explanation of the mechanism, the spatial distribution of each type of facilities and the per capita level are important parameters to reflect the level of city public facilities; the distribution density ensures that the people could find facilities in a certain traffic distance. The per capita amount of ownership avoids the consumption of public goods caused by excessive consumption groups.

In the detailed analysis, the number of public facilities in the grid is summarized and the facility density D_i is calculated. That is, $D_i = \sum N_{ik}/S_i$ ($i \in j, S_i \neq 0$), where S_i is the area of the grid i , j is the collection of grid i , and the spatial distribution of the facility density is analyzed by means of kernel density.

The spatial distribution of public facilities density in Shenzhen city is shown in Figure 4. The high value mainly concentrated in Futian–Luohu Stretching region, which sits near Hong Kong. The facilities density of some parts of Yuehai, Longgang, Xixiang, and Sinan sub-region is higher, while that of the Guangming, Yantian, Pingshan, Dapeng, and Longhua District is not high,

which shows that the facilities supply level there is relatively low.

After the completion of the overall facility analysis, the number of every type of facilities within the grid is summarized respectively. Similarly, as calculated above, define the distribution density of each facilities $D_{ik} = \sum N_{ik}/S_i$ ($i \in j, S_i \neq 0$), where S_i is the area of the grid i , N_{ik} is the number of facilities belonged to type k of the grid i . The density distribution of the facility is obtained by the kernel density method.

The density distribution of all kinds of facilities is shown in Table 1. For example, supermarkets and other convenience shopping facilities are widely distributed in Baoan District, thereby the supermarket density in Xin'an and Xixiang sub-region is the highest, followed by Huangbei Road in Luohu District and Shatou street in Nanshan District. Density distribution of facilities is no longer described due to the limitation of article length.

On the basis above, we overlay resident population in the construction land and the number of facilities within the grid, and then analyze spatial distribution of the amount of per capita ownership of facilities, in which the construction land data are the aerial one in 2014; the resident population comes from the sixth census data. We calculate the number of grid C_i construction land occupied by each town, and then calculate the grid resident population:

$$P_i = \frac{T_j}{\sum C_i} \quad (1)$$

where T_i indicates that the number of resident population of the town, $\sum C_i$ indicates the number of grid of construction land, and then calculate the amount of each grid's per capita facilities.

$$M_i = \frac{\sum N_i}{P_i} \quad (2)$$

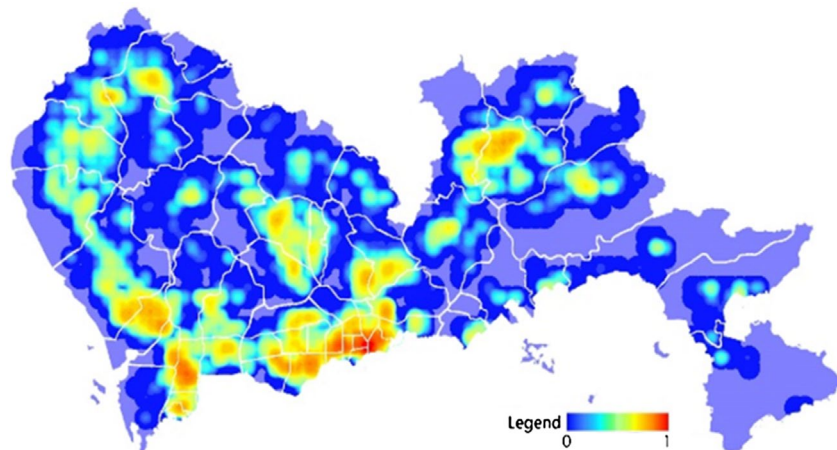
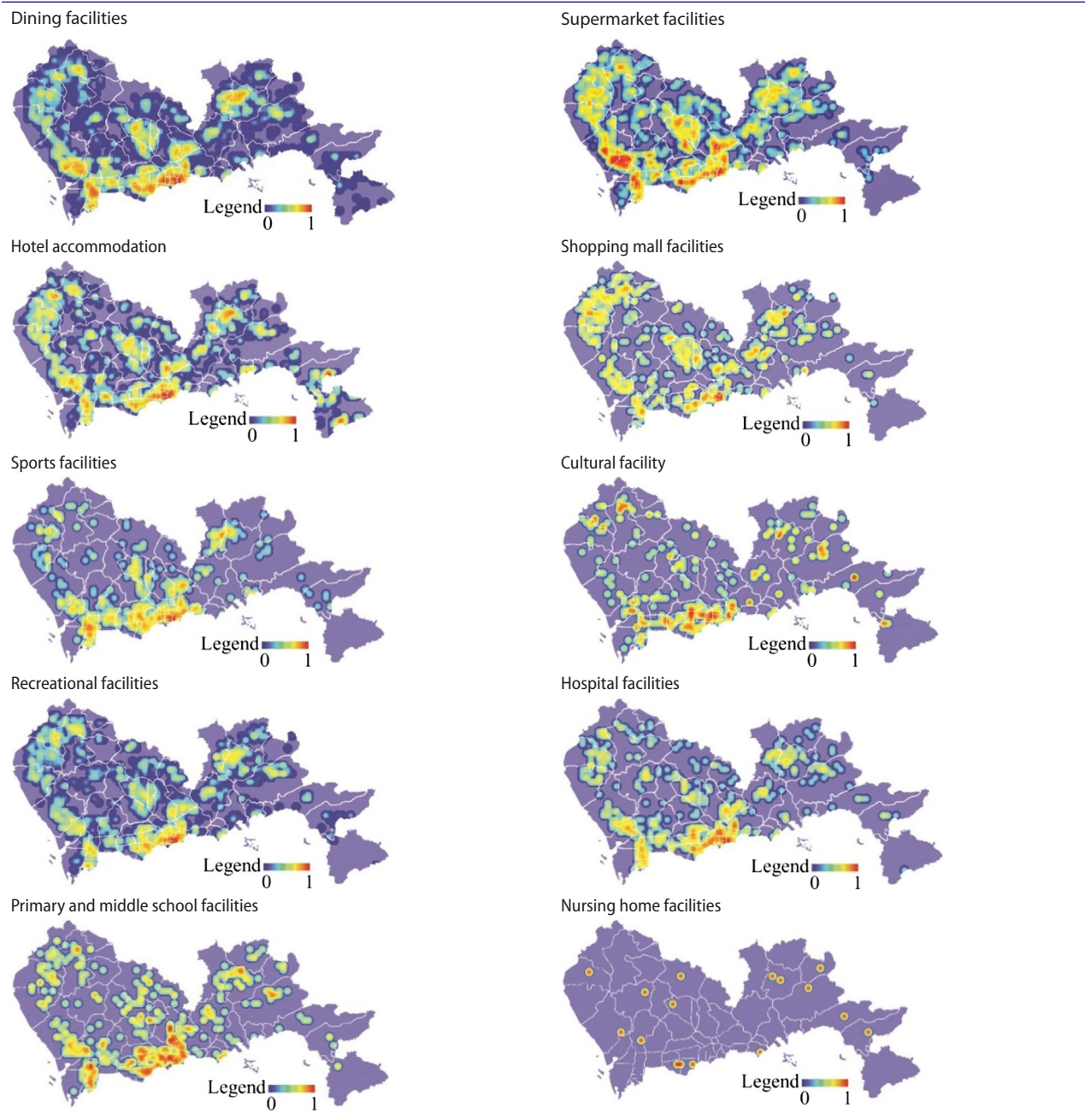


Figure 4. Spatial distribution of public service facility density.

Table 1. Kernel density distribution of public facility.



where $\sum N_i$ represents the total number of facilities in the grid. On this basis, the spatial distribution of per capita amount of facilities is expressed by means of nuclear density analysis.

As shown in Figure 5, after synthesizing all kinds of public facilities, we find that the area of highest per capita facilities amount is Longgang sub-region and the second one is Kuichong. The amount of per capita facilities ownership in Yellow Bay and Dongmen in Luohu District is also high, but the rest needs to be improved. On the one hand, it reflects that the density of facilities in peripheral areas with sparse population basically meets the local demand. On the other hand, with the increase in the population in the central area, the phenomenon of per capita possession declining and service overloading is bound to exist.

5. Subjective evaluation of public facilities

The subjective evaluation includes the evaluation of the attention and satisfaction degree of each facility. For a facility, the degree of attention (the number of evaluations) reflects the importance of the facility. And in the subjective evaluation, according to the relevant research experience at home and abroad, generally using structured questionnaire method of Likert scales, five files or seven files for evaluation objects are evaluated. Considering that the facilities data in www.Dianping.com include the number of comments and evaluation score, the study will take the data as a basic option for subjective evaluation.

In detail, the evaluation of all kinds of facilities in the grid will be aggregated to get the attention amount of

facilities in the grid. For evaluation scores, according to the impact of the above 10 types of facilities to urban public facilities, based on the review of Internet users on the 5-point system for each facility, the weight values of all kinds of facilities are obtained by using Phil's method and analytic hierarchy process (AHP). The facility, rated higher than the average value of a standard deviation, is regarded as excellent facility. Then, we obtain the optimal spatial distribution of these facilities in Shenzhen. Finally, we normalize the optimal score and intercept along three traffic sections: Guangshen highway to Shenzhen West Railway Station of Subway Line 1, Shenzhen West Railway Station of Subway Line 1 to Luosa Road, and Shuanglong Railway Station of Subway Line 3 to Shenzhen Railway Station. We then analyze the satisfaction distribution of facilities in the cross section of the main passenger traffic.

For a certain kind of facilities, this article chose the way of the mean-square deviation to determine the weights of individual indicators. The basic principle is to calculate variance to all the variables of each type of facilities and to use the result as the weight. The bigger the spatial difference, the higher the weight, and vice versa. The study concentrates on the variables that are larger in the spatial distribution. More generally, the

improvement of space benefits is meaningful only if it does not harm the interests of others, which is the basic viewpoint of Pareto theory. Therefore, the weight of the average variance of the individual facilities is calculated by mean-square deviation method. In detail, the weight of all kinds of facilities score is calculated, and the weight of k in the system is calculated by ω_k .

$$\omega_k = \frac{\sigma_k}{\sum_{k=1}^{10} \sigma_k} \quad (3)$$

$$\text{where } \sigma_k^2 = \frac{\sum_{i=1}^n (x_{ik} - \bar{x}_k)^2}{n}, \bar{x}_k = \frac{\sum_{i=1}^n x_{ik}}{n}.$$

As shown in Figure 6, the highest degree of concern of regional facilities is Dongmen, Huangbei and Nanhu sub-region belonged to Luohu District, followed by Yuanling in Futian District, Yuehai in Nanshan District. In addition, the facility of Guangming New District and the Longgang District also holds a high degree of concern, the other districts are relatively close with no obvious peak area.

From respective preferential score (satisfaction score) distribution (Figure 7), facilities satisfaction in Longgang and Longcheng is higher compared with that in other areas, but the overall level of satisfaction does not present the same. There is some peak satisfaction

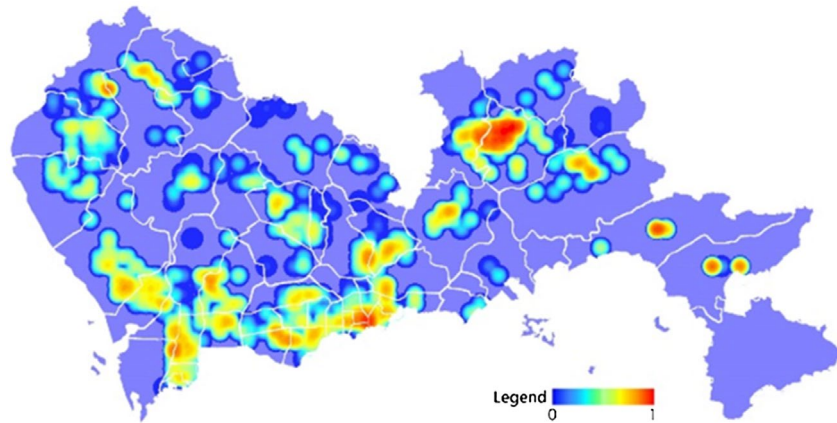


Figure 5. Spatial distribution of public services facility per capita.

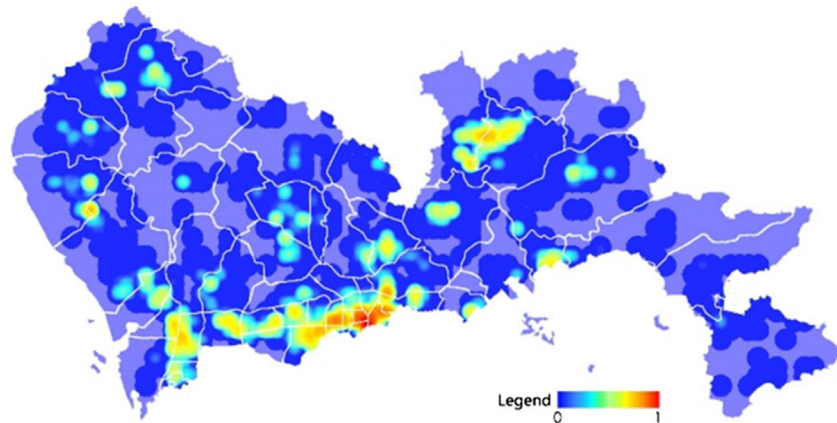


Figure 6. Spatial distribution of attention to public service facility.

score that appeared in the area of Futian, Dongmen, and Guangming New District, and the higher level of satisfaction also took place in Luohu District and Futian District. Multi-point spatial distribution of high satisfaction evaluation score to a certain extent matches the polycentric structure of Shenzhen.

Further, by means of superposition facility satisfaction with main traffic routes, it is found that the spatial distribution of the excellent evaluation facilities is closely related to the public transport sites in the three traffic routes, as shown in Figures 8–10. Among them, the service facilities satisfaction level is higher around Baisong interchange in the view of the section from Guangshen highway to Shenzhen West Railway Station of Subway Line 1. Furthermore, the overall satisfaction of all kinds of public facilities is relatively high around the center of Shenzhen West Railway Station, and the radiation range is big. On the east of Shenzhen West Railway Station of Subway Line 1, the main site of the facilities satisfaction is high, among which China World Trade Center station is in the first place and Convention and Exhibition Center and Luohu Station are in second and third, respectively. It is worth noting that, Window of World, built in 1994, with the rise of time and emerging markets, the level of satisfaction of the facilities is declining. Facilities satisfaction is relatively high in the vicinity of Longcheng Square

Station of Subway Line 3, followed by Luohu Station and Shenzhen Station. Overall, facilities satisfaction evaluation distribution appears to have some uplift of the peak area, directly reflecting the polycentric space structure of Shenzhen city.

6. Comprehensive evaluation of public facilities

For comprehensive evaluation, the research overlaps the spatial data of the subjective and objective evaluation and the weight of the subjective/objective evaluation of the public facilities is calculated, finally forming the comprehensive evaluation of urban public facilities evaluation. In specific calculation, we first get all kinds of sub-index value. Then, we determine the index weight according to the variance of the value in the basic elements of each facility type. Third, the study determines the comprehensive weight combining with the variance and Delphi method. Finally, 10 evaluation indexes are thus obtained.

In the basic model of the comprehensive evaluation of the research, the subjective/objective evaluation of individual facilities is firstly calculated:

$$y_{ik} = \sum_{s=1}^4 \omega_{ks} x_{iks} \quad (4)$$

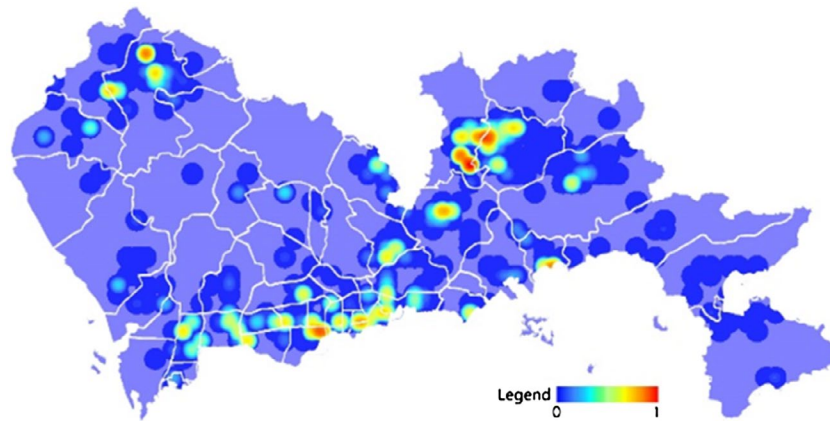


Figure 7. Spatial distribution of satisfaction to public service facility.

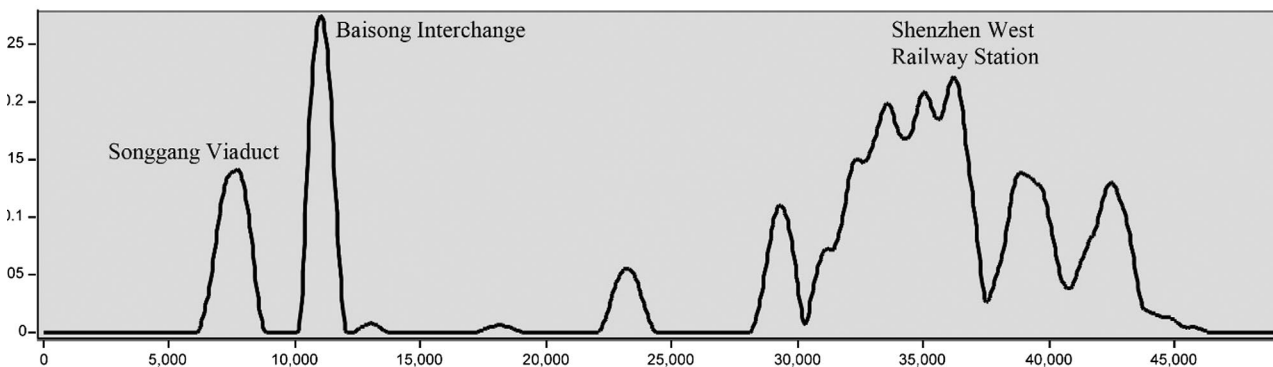


Figure 8. Distribution of public facility satisfaction from Guangshen highway to Shenzhen West Station of Subway Line 1.

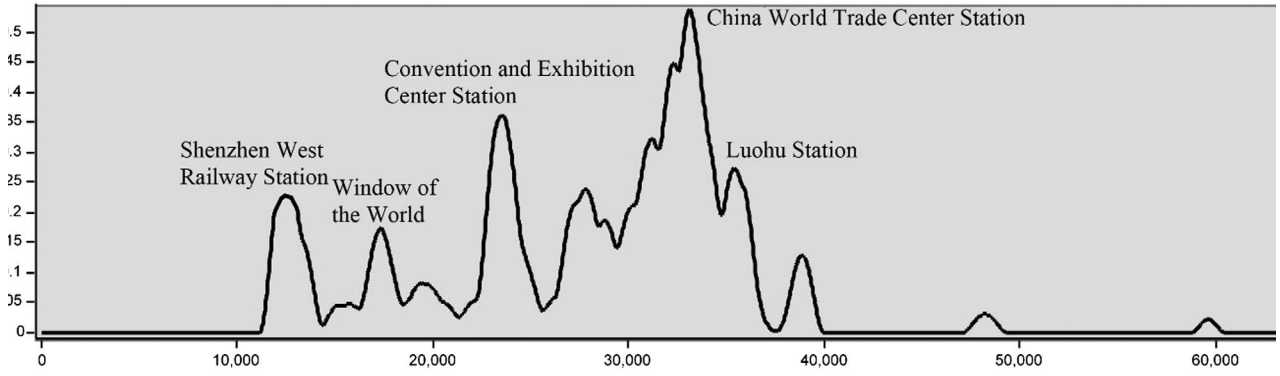


Figure 9. Distribution of public facility satisfaction from Shenzhen West Railway Station of Subway Line 1 to Luosha Road.

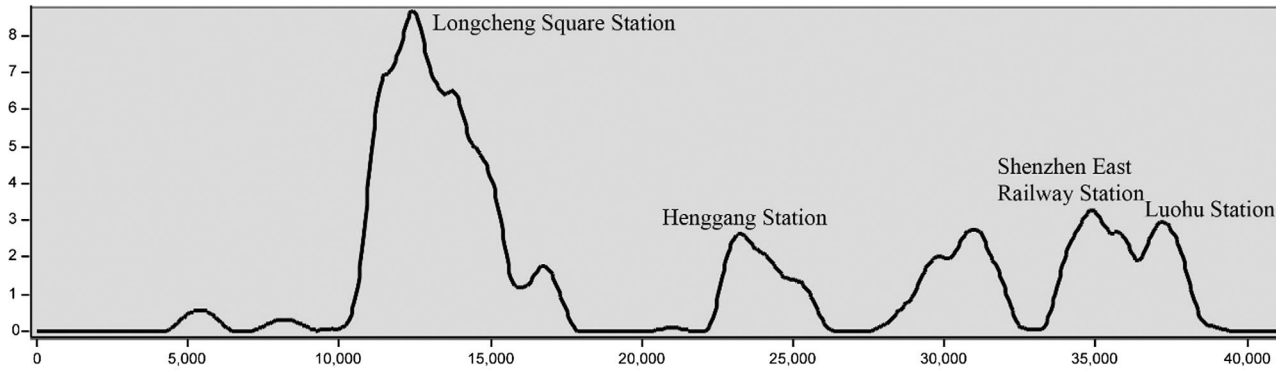


Figure 10. Distribution of satisfaction to public facility under Subway Line 3 (from Shuanglong Station to Shenzhen Railway Station).

where x_{iks} represents the score of space grid i of facility k in sub-index s , among which $i = 1, 2, \dots, n$; facility $k = 1, 2, \dots, 10$; sub-index $s = 1, 2, 3, 4$. And then the weight between the facilities is set up:

$$y_i = \sum_{k=1}^{10} \beta_k y_{ik} \quad (5)$$

Aiming at the problem of setting different weights between facilities, based on the characteristic that utility function represents the addition and subtraction relationship after logarithm processing, Delphi importance ranking method is increased in comprehensive evaluation, and applied into final comprehensive assignment along with weights. In detail, according to Hong Yan's comprehensive index weight calculation process (Hong 2008), first, we calculate the arithmetic mean value of secondary indicators \bar{x}_{ks} , then calculate the average of them σ_{ks}^2 and the weight ω_{ks} , finally get the score y_{ik} of the grid i of the facility k . Ten kinds of facilities evaluation values thus obtained are further applied to determine the comprehensive weights based on variance and Delphi method. According to the Delphi weighting method, we first determine the evaluation of k facilities weight p_k . According to the standard deviation method, \bar{y}_k, σ_k^2 were calculated to determine the weight of each facility, where \bar{y}_k is the mean value of y_{ik} , representing the arithmetic mean value of the k facility, and the σ_k^2 represents the standard deviation of the facility k . According to

Lagrange's method, c_1, c_2 are calculated, respectively, and the standard coefficients c_1^*, c_2^* of p_k and q_k are calculated, respectively. To find the weight of all kinds of facilities β_k , we use the integrated method, $\beta_k = c_1^* p_k + c_2^* q_k$. The comprehensive evaluation of the public facilities of the space grid unit i is:

$$y_i = \sum_{k=1}^{10} \beta_k y_{ik} = \sum_{k=1}^{10} (c_1^* p_k + c_2^* q_k) y_{ik} \quad (6)$$

Research attributes weight assignment on the classification of all kinds of facilities and on the subjective/objective evaluation index for each category of facilities, so the weight is directly related to the final evaluation results. In the subjective and objective comprehensive evaluation, for each kind of facilities, ω_{ks} is used to evaluate the weights of facility k and secondary index s in the evaluation system. This value is calculated by mean-square deviation method.

$$\omega_{ks} = \frac{\sigma_{ks}}{\sum_{s=1}^t \sigma_{ks}} \quad (7)$$

$$\text{where } \sigma_{ks}^2 = \frac{\sum_{i=1}^n (x_{iks} - \bar{x}_{ks})^2}{n}, \bar{x}_{ks} = \frac{\sum_{i=1}^n x_{iks}}{n}.$$

β_k is the weight of facilities and the integrated weight method is used in this study. p_k and q_k , respectively, are weighted by Delphi method and mean-square deviation method.

$$\beta_k = c_1^* p_k + c_2^* q_k \quad (8)$$

$$\text{where } c_1^* = \frac{c_1}{c_1 + c_2}, c_2^* = \frac{c_2}{c_1 + c_2}.$$

And $c_1 > 0, c_2 > 0$, thus the comprehensive evaluation of the public facilities of the space grid unit i is:

$$y_i = \sum_{k=1}^{10} \beta_k y_{ik} = \sum_{k=1}^{10} (c_1^* p_k + c_2^* q_k) y_{ik} \quad (9)$$

According to Lagrange's method:

$$c_1 = \frac{\sum_{i=1}^n \sum_{k=1}^{10} p_k y_{ik}}{\sqrt{(\sum_{i=1}^n \sum_{k=1}^4 p_k y_{ik})^2 + (\sum_{i=1}^n \sum_{k=1}^4 q_k y_{ik})^2}},$$

$$c_2 = \frac{\sum_{i=1}^n \sum_{k=1}^{10} q_k y_{ik}}{\sqrt{(\sum_{i=1}^n \sum_{k=1}^4 p_k y_{ik})^2 + (\sum_{i=1}^n \sum_{k=1}^4 q_k y_{ik})^2}} \quad (10)$$

Here, p_k is the weight of the facility k evaluation by AHP method, which could be determined by order relation method. q_k means the weight of the facility k evaluation

by objective mean-square deviation weighting method. At the same time, \bar{y}_k is the mean of y_{ik} . We use the following formula to determine the value of q_k .

$$q_k = \frac{\sigma_k}{\sum_{s=1}^t \sigma_s} \quad (11)$$

$$\text{where } \sigma_k^2 = \frac{\sum_{i=1}^n (y_{ik} - \bar{y}_k)^2}{n}.$$

Based on this method, we eventually form the comprehensive assessment of subjective/objective evaluation of 10 public facilities in Shenzhen, and the final results are shown in Figure 11. The higher score area of comprehensive evaluation concentrates in Luohu District and Futian District, and a continuous region appears from Luohu Station as the center along the Subway Line 1. At the same time, because of the advantages of per capita ownership, Longgang District has a high score of comprehensive evaluation, and has a high degree of satisfaction. The rest of the districts only partially have some high degree area of evaluation, which needs to be improved. The facility level in Nanshan District and Baoan District, along with the Guangzhou-Shenzhen highway and Subway Line 1, has potential improvement.

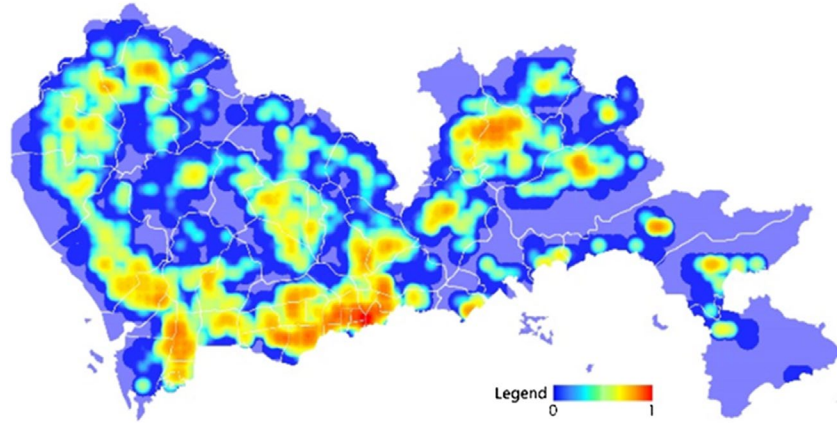


Figure 11. Spatial distribution of comprehensive evaluation.

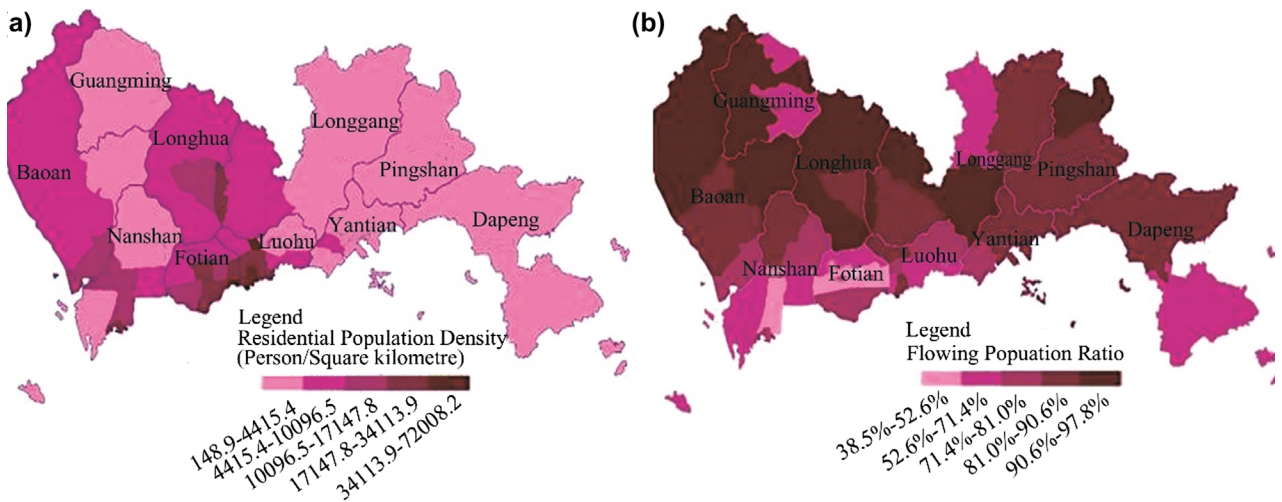


Figure 12. Spatial distribution of resident population density (a) and flowing population proportion (b).

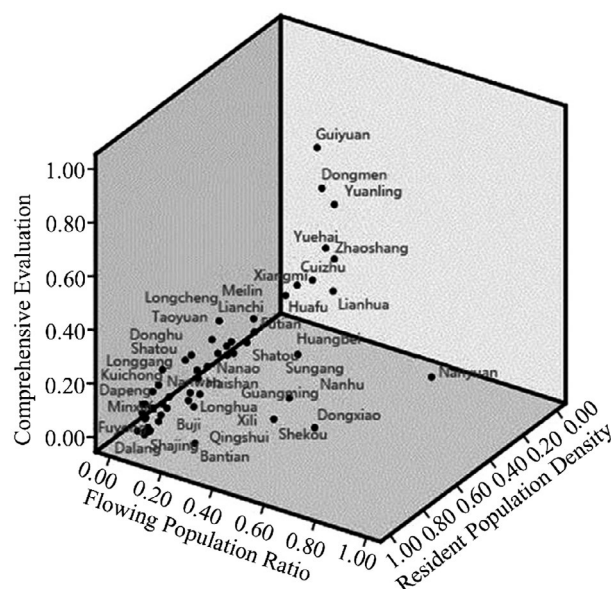


Figure 13. Comprehensive evaluation with per capita in each zone.

7. Discussion

This study uses data crawling approach to get access to network evaluation data of public facilities. Then through address resolution, spatial dimension reduction, and other data conversion, from the structural level, the research realizes the evaluation of urban public facilities and reflects the social micro individual satisfaction, thereby achieving the measurement of public facilities in Shenzhen. According to the scaling law (Bettencourt et al. 2007; Batty 2008; Bettencourt and West 2010), both population size and its spatial distribution influence the various elements of city, and the negative effect caused by excessive agglomeration becomes increasingly apparent with scale expansion.

In addition, it is worthy of pointing out that there are lots of foreign population in Shenzhen. In this study, the distribution density of the facilities and the distribution characteristics of the population were compared, and the distribution of public facilities and floating population in Shenzhen is revealed. According to the data of the sixth census of demographics, the high proportion of the foreign population is mainly concentrated in Baoan District, Guangming New District and Longhua District (Figure 12). From the public facility evaluation, resident population density, and the standard value of foreign population distribution (Figure 13), there is no obvious peak area of the spatial quality compared to high quality area such as Dongmen and Guiyuan in Luohu District. There is no peak area with high evaluation in the external population agglomeration region, such as Fuyong and Songgang in Baoan District, Dalang in Longhua District, Bantian in Longgang District, etc. This phenomenon shows that the social welfare in Shenzhen has the problem of partially spatial imbalance. It also indicates that the social reality of urban space is truly

reflected by the new data. Furthermore, it reflects the rationality of the results of the empirical analysis.

Limited by the fact that different social groups concerned about different facilities object, in the urban-scale evaluation of the research, the results of the study inevitably lead to bias, and this article reflects the group of the users of www.dianping.com. Therefore, the follow-up study is more important. Finally, it is shown that the effectiveness of open source data in planning practice is mainly reflected in the continuous improvement of the tool rationality, but the obtaining of new technology and new resources cannot replace the value of the planning policy. Planners, in the diverse needs of network society, will face more complex value judgment.

Funding

This work was funded by National Natural Science Foundation of China [grant numbers. 51478189 and 51308220]; Natural Science Foundation of Guangdong [grant number 2014A03031326]; and Fundamental Research Funds for the Central Universities [grant number 2015ZZ0022].

Notes on contributors

Miaoxi Zhao is Associate Professor of School of Architecture in South China University of Technology and also Visiting Professor of Geography Department in Ghent University. His research interests include urban science, regional planning, and urban network research.

Gaofeng Xu is a graduate student of School of Architecture in South China University of Technology. His research interests include regional planning and urban planning management.

Yun Li is Associate Professor of School of Architecture and Urban Planning in Shenzhen University. His research focuses on urban planning and urban design.

References

- Batty, M. 2008. "The Size, Scale, and Shape of Cities." *Science* 319 (5864): 769–771. doi:<http://dx.doi.org/10.1126/science.1151419>.
- Batty, M. 2013. "Big Data, Smart Cities and City Planning." *Dialogues in Human Geography* 3 (3): 274–279. doi:<http://dx.doi.org/10.1177/2043820613513390>.
- Becker, R. A., R. Caceres, K. Hanson, and J. M. Loh. 2011. "A Tale of One City: Using Cellular Network Data for Urban Planning." *IEEE Pervasive Computing* 10 (4): 18–26. doi:<http://dx.doi.org/10.1109/MPRV.2011.44>.
- Bettencourt, L. M. A., J. Lobo, D. Helbing, C. Kühnert, and G. B. West. 2007. "Growth, Innovation, Scaling, and the pace of Life in Cities." *Science* 104 (17): 7301–7306. doi:<http://dx.doi.org/10.1073/pnas.0610172104>.
- Bettencourt, L., and G. West. 2010. "A Unified Theory of Urban Living." *Nature* 467 (7318): 912–913. doi:<http://dx.doi.org/10.1038/467912a>.
- Castells, M. 1996. *The Rise of the Network Society*. Oxford: Blackwell.
- Chen, L. 2012. "An Evaluation of Public Facility Based on Network Data: A Case Study on Hospital Facility in Lankao Town, Henan Province." The Summary of the 7th National

- Symposium on Geography, Henan, October China, 11–12, 42 (in Chinese).
- Cranshaw, J., R. Schwartz, J. I. Hong, and N. Sadeh. 2012. "The Livelihoods Project: Utilizing Social Media to Understand the Dynamics of a City." In *Proceedings of the 6th International AAAI Conference on Weblogs and Social Media (ICWSM-12)*, Dublin, Ireland, 4–7.
- Delyser, D., and D. Sui. 2013. "Crossing the Qualitative-Quantitative Divide II: Inventive Approaches to Big Data, Mobile Methods, and Rhythmanalysis." *Progress in Human Geography* 37 (2): 293–305. doi:<http://dx.doi.org/10.1177/0309132512444063>.
- Gu, M. D., and H. W. Yin. 2010. "Review on Spatial Accessibility and Equity of Public Facilities." *Urban Problems* 5: 25–29. doi:<http://dx.doi.org/10.13239/j.bjsshkxy.cswt.2010.05.008>. (in Chinese).
- Gui, Z. M., Y. Xiang, and Y. J. Li. 2012. "Parallel Discovering of City Hot Spot Based on Taxi Trajectories." *Journal of Huazhong University of Science & Technology (Natural Science Edition)* 40: 187–190. doi:<http://dx.doi.org/10.13245/j.hust.2012.s1.045>. (in Chinese).
- Han, Y. H., and Y. Q. Lu. 2012. "Accessibility Assessment and Planning of Public Facilities for Education: A Case Study on Senior High Schools in Yizheng City." *Scientia Geographica Sinica* 32 (7): 822–827. doi:<http://dx.doi.org/10.13249/j.cnki.sgs.2012.07.007>. (in Chinese).
- Hollenstein, L., and R. Purves. 2010. "Exploring Place through User-Generated Content: Using Flickr tags to Describe City Cores." *Journal of Spatial Information Science* 1: 21–48. doi:<http://dx.doi.org/10.5311/JOSIS.2010.1.3>.
- Hong, Y. 2008. "The Comprehensive Evaluation of the Coordinative Development in Northeastern Provinces of China." Master's Thesis, Northeastern University. (in Chinese).
- Hu, Y. Z., and J. Wu. 2011. "Chinese Cities Have a Huge Challenge to Maintain a Decent Quality of Life for Their Citizens: Dialogue with Sir Peter Hall." *Urban Planning International* 26 (3): 111–114. doi:<http://dx.doi.org/10.3969/j.issn.1673-9493.2011.03.017>. (in Chinese).
- Kang, C. G., Y. Zhang, X. J. Ma, and Y. Liu. 2013. "Inferring Properties and Revealing Geographical Impacts of Intercity Mobile Communication Network of China Using a Subnet Data Set." *International Journal of Geographical Information Science* 27 (3): 1–18. doi:<http://dx.doi.org/10.1080/13658816.2012.689838>.
- Kitchin, R. 2013. "Big Data and Human Geography: Opportunities, Challenges and Risks." *Dialogues in Human Geography* 3 (3): 262–267. doi:<http://dx.doi.org/10.1177/2043820613513388>.
- Liu, J. N., Y. Fang, and K. Guo. 2014. "Research Progress in Location Big Data Analysis and Processing." *Geomatics and Information Science of Wuhan University* 4: 379–385. doi:<http://dx.doi.org/10.13203/j.whugis20140210>. (in Chinese).
- Long, Y., Y. Zhang, and C. Y. Cui. 2012. "Identifying Commuting Pattern of Beijing Using Bus Smart Card Data." *Acta Geographica Sinica* 67 (10): 1339–1352. doi:<http://dx.doi.org/10.11821/xb201210005>. (in Chinese).
- Lüscher, P., and R. Weibel. 2013. "Exploiting Empirical Knowledge for Automatic Delineation of City Centres from Large-Scale Topographic Databases." *Computers, Environment and Urban Systems* 37: 18–34. doi:<http://dx.doi.org/10.1016/j.compenvurbsys.2012.07.001>.
- Mark, B., and M. Nick. 2011. "Microscopic Simulations of Complex Metropolitan Dynamics." *Amsterdam: Complex City*. Accessed March 25, 2016. http://eprints.ncrm.ac.uk/2051/1/complex_city_paper%5B1%5D.pdf
- Naaman, M., A. X. Zhang, and S. Brody. 2012. "On the Study of Diurnal Urban Routines on Twitter." In *Proceedings of Sixth International AAAI Conference on Weblogs and Social Media*, Dublin, Ireland, June, 2012, 258–265.
- Niu, X. Y., L. Ding, and X. D. Song. 2014. "Understanding Urban Spatial Structure of Shanghai Central City Based on Mobile Phone Data." *Urban Planning Forum* 6: 61–67. doi:<http://dx.doi.org/10.3969/j.issn.1000-3363.2014.06.009>. (in Chinese).
- Qin, X., F. Zhen, and L. F. Xiong. 2013. "Methods in Urban Temporal and Spatial Behavior Research in the Big Data Era." *Progress in Geography* 32 (9): 1352–1361. doi:<http://dx.doi.org/10.11820/dlkxjz.2013.09.005>. (in Chinese).
- Ren, Z. C. 2014. "Assessment of the Status Quo of Public Facilities Allocation Based on the User – A Case Study of Tourist Oriented Small Towns." In *Proceedings of 2014 Annual Meeting of Chinese Urban Planning*, Haikou, China, September, 2014, 13–15. (in Chinese).
- Sun, Y., K. Y. Wang, and X. D. Yao. 2015. "Economic Benefits Evaluation of Urban Public Infrastructure Based on the DEA Cross-Efficiency Method." *China Soft Science* 1: 172–183. doi:<http://dx.doi.org/10.3969/j.issn.1002-9753.2015.01.018>. (in Chinese).
- Wang, D., W. Zhu, and D. C. Xie. 2014. "The Analysis Framework, Difficulties and Existing Progress of Urban Space Structure Based on the Signaling Data of Mobile Phone." Accessed March 25, 2016. <http://bbs.caup.net/read-htm-tid-30130-page-1.html>
- Yang, J. Y., Y. Shi, and D. R. Deng. 2010. "The Suitable Layout of Public Facilities: Nanjing's Binjiang New City." *Planners* 26 (4): 19–24. doi:<http://dx.doi.org/10.3969/j.issn.1006-0022.2010.04.003>. (in Chinese).
- Yang, Q. Y., and J. G. Xu. 2015. "An Analysis Model of Public Facilities Based on Attraction Accessibility." *Planners* 31 (7): 96–101. doi:<http://dx.doi.org/10.3969/j.issn.1006-0022.2015.07.016>. (in Chinese).
- Yu, W. H., and T. H. Ai. 2015. "The Visualization and Analysis of POI Features under Network Space Supported by Kernel Density Estimation." *Acta Geodaetica Et Cartographica Sinica* 44 (1): 82–90. doi:<http://dx.doi.org/10.11947/j.AGCS.2015.20130538>. (in Chinese).
- Zhang, X. X. 2014. "Spatial Analysis in the Era of Big Data." *Geomatics and Information Science of Wuhan University* 39 (6): 655–659. doi:<http://dx.doi.org/10.13203/j.whugis20140143>. (in Chinese).
- Zhang, Y. Z., T. H. Wei, and C. Hua. 2012. "Research on the Spatial Equity in the Distribution of Public Facilities Based Section Quotient." *Huazhong Architecture* 14 (2): 38–40. doi:<http://dx.doi.org/10.3969/j.issn.1003-739X.2012.02.010>. (in Chinese).
- Zhao, M. X. 2009. "Study on Population Age Structure and Community Service Facilities Planning and Allocation Index – A Case Study of Jiangwan City Shanghai and Liangcheng Community." *Journal of Nanjing College for Population Programme Management* 1: 25–28. doi:<http://dx.doi.org/10.3969/j.issn.1007-032X.2009.01.005>. (in Chinese).